Available online at www.sciencedirect.com
SCIENCE DIRECT®
Remote Sensing of Environment 102 (2006) 12–23
Remote Sensing of Environment
www.elsevier.com/locate/rse

A multi-sensor approach for the on-orbit validation of ocean color satellite data products

Sean W. Bailey [a,*], P. Jeremy Werdell [b]

[a] Futuretech Corporation, Greenbelt, MD, USA
[b] Science Systems and Applications Inc., Lanham, MD, USA

Received 4 November 2005; received in revised form 19 January 2006; accepted 19 January 2006

Abstract

The validation of satellite ocean color data products is a critical component in establishing their measurement uncertainties, assessing their scientific utility, and identifying conditions for which their reliability is suspect. Such efforts require a considerable amount of high quality in situ data, preferably consistently processed and spanning the satellite mission lifetime. This paper outlines the NASA Ocean Biology Processing Group's (OBPG) method for validating satellite data products using in situ measurements as ground truth. Currently, the OBPG uses the described method for validating several historical and operational ocean color missions. By way of a case study, results for the Sea-Viewing Wide Field-of-View Sensor (SeaWiFS) are shown. These results indicate that for the majority of the global ocean, SeaWiFS data approach the target uncertainties of ±5% for clear water radiances as defined prior to launch. Our results add confidence in the use of these data for global climate studies, where a consistent, high quality data set covering a multi-year time span is essential.

PACS: 42.68.Xy; 92.70.j; 92.70.Jw
Keywords: Ocean color; Satellite validation; SeaWiFS; Remote sensing; Water-leaving radiance; Chlorophyll

1. Introduction

Estimating the rates and magnitudes of ocean primary productivity on regional and global scales is key to understanding the role of the ocean in the Earth's carbon cycles (Behrenfeld & Falkowski, 1997b; Longhurst et al., 1995; Kuring et al., 1990; Prasad & Haedrich, 1994). The synoptic views of the marine biosphere captured by satellite-based ocean color instruments provide valuable data at spatial and temporal scales unattainable with shipboard or moored instrumentation. This was aptly demonstrated by the proof-of-concept Coastal Zone Color Scanner (CZCS) (Gordon et al., 1983; Hovis et al., 1980). Drawing on its successful legacy, a number of advanced ocean color satellite instruments were launched in the past decade [e.g., the Ocean Color and Temperature Scanner–OCTS (Iwasaki et al., 1992), the Sea-viewing Wide Field-of-view Sensor–SeaWiFS (Hooker et al., 1992), the Moderate Resolution Imaging Spectroradiometer–MODIS (Salomonson et al., 1989), and the Medium Resolution Imaging Spectrometer–MERIS (Rast & Bezy, 1999)]. More are scheduled for launch in the near future [e.g., the Visible Infrared Imager/Radiometer Suite–VIIRS (Welsch et al., 2001)–on board the National Polar-orbiting Operational Environmental Satellite System].

Space-borne ocean color instruments measure the spectrum of sunlight reflected from ocean waters at selected visible and near-infrared wavebands. These radiance spectra are used to estimate geophysical parameters, such as the surface concentration of the phytoplankton pigment chlorophyll *a*, C_a, via the application of bio-optical algorithms (O'Reilly et al., 1998). These derived data products are subsequently input into secondary (i.e., higher order) geophysical algorithms, as is the case for marine primary production, PP (e.g., Behrenfeld and Falkowski, 1997a). The uncertainties in global estimates of C_a and PP are contingent on the uncertainties of their model input parameters (e.g., spectral reflectance and C_a). Predictably, uncertainty increases for secondary algorithms that require derived products as input (Behrenfeld & Falkowski, 1997b).

* Corresponding author.
E-mail address: sean.w.bailey.1@gsfc.nasa.gov (S.W. Bailey).

0034-4257/$ - see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.rse.2006.01.015

Global accuracy goals for spectral reflectance and C_a for modern sensors are commonly defined as 5% and 35%, respectively, in clear, natural waters (Hooker et al., 1992). As such, statistical validation of these products is prerequisite in verifying that such goals are being met. Following McClain et al. (2002), we define validation as "the process of determining the spatial and temporal error fields of a given biological or geophysical data product". The NASA Ocean Biology Processing Group (OBPG) at the Goddard Space Flight Center executes satellite validation activities via the direct comparison of remotely sensed measurements with coincident in situ measurements. The OBPG maintains responsibility for the operational processing of ocean color data within NASA, as well as the post-launch calibration, validation, and subsequent distribution of the data products. A significant component of this responsibility is quantifying how well the satellite-retrieved products reflect true conditions. As will be highlighted in a subsequent section, a comprehensive in situ data set with measurements covering a wide range of oceanographic conditions is essential in this process (Werdell & Bailey, 2005).

Much recent refereed research on ocean color validation efforts focuses on comparisons of satellite-derived C_a retrievals with regional in situ data sets, many of which are based on specific, or single, field campaigns (e.g. Barbini et al., 2005; D'Ortenzio et al., 2002; Gohin et al., 2002; He et al., 2000; Smyth et al., 2002). While a necessary step towards a comprehensive understanding of the uncertainties in global primary production models, validation of satellite-retrieved C_a alone is insufficient, as uncertainties in the retrievals are strongly affected by the uncertainties in the C_a algorithm applied. Moreover, as many C_a algorithms make use of reflectance ratios (O'Reilly et al., 1998), underlying problems with radiometry are often masked. Since the primary measurement of satellite-based sensors is spectral radiance, the focus of the OBPG validation activity is on the retrieved water-leaving radiance ($L_{wn}(\lambda)$) estimates. Several other independent studies include validation of spectral reflectance (e.g., Froidefond et al., 2002; Gordon et al., 1983; Pinkerton & Aiken, 1999); however, these also suffer from the limitations imposed by short-term, regional or cruise-specific data sets.

In developing their calibration and validation program, the OBPG adopted the approach of globally validating satellite ocean color sensors throughout the life of their missions, particularly for data products to be used in long-term multi-sensor time series (Barnes et al., 2003; Donlon et al., 2002). Without continuous validation for monitoring the long-term stability of satellite instruments, such efforts are hindered as uncertainties in sensor calibration are not tracked, and potential instrument effects may, therefore, be misinterpreted as geophysical phenomena. This is particularly true for remotely sensed ocean color as the $L_{wn}(\lambda)$ component accounts for only about 10% of the total reflectance signal received by the sensor at the top of the atmosphere in the visible wavelengths (Gordon, 1998).

Ultimately, experience has demonstrated that a number of benefits are realized with a rigorous validation activity, including: (i) the assignment of a measure of accuracy to satellite-derived products, which lends confidence to their scientific utility in higher-order derived products (Behrenfeld & Falkowski, 1997b); (ii) the verification of on-orbit satellite calibration (Barnes et al., 2001); (iii) the evaluation of the long-term stability of satellite measurements (Franz et al., 2005); and (iv) the identification of conditions, either oceanic, atmospheric or satellite specific, for which satellite-derived products are invalid. In this paper, we outline and define the NASA OBPG satellite validation approach. The method described may be applied to most sensors with few modifications, and in fact, the OBPG incorporates these methods in their current OCTS, SeaWiFS, and MODIS validation activities. Results for SeaWiFS are presented.

2. Methods

2.1. In situ data

The OBPG maintains a local repository of in situ bio-optical data, the SeaWiFS Bio-optical Archive and Storage System (SeaBASS) (Hooker et al., 1994; Werdell et al., 2003), with the purpose of acquiring a data set of sufficient size, quality, and diversity to support and sustain its regular scientific analyses. SeaBASS is populated with both voluntary and funded data contributions from investigators worldwide (Fargion et al., 2004). To develop consistency across multiple contributors and institutions, the NASA Sensor Inter-comparison and Merger for Biological and Interdisciplinary Oceanic Studies (SIMBIOS) Project (McClain et al., 2002) defined and documented a series of in situ sampling strategies and data requirements that ensure that the holdings of SeaBASS are appropriate for both algorithm development and ocean color sensor validation (Mueller et al., 2003a,b). To further ensure consistency, the radiometric and pigment data within SeaBASS were recently re-evaluated to produce remote-sensing relevant surface values using a common set of processing and quality control steps (Werdell & Bailey, 2005). These latter data are incorporated into the OBPG satellite validation activity.

While in situ measurements are sometimes referred to as 'ground-truth' measurements, they are rarely 'absolute truth'. Full characterization of the inherent measurement error of the field instrument is essential for any validation effort. Ideally, the instruments are also well calibrated and deployed in a manner consistent with well-defined protocols (Mueller et al., 2003a,b). To reduce systematic uncertainties resulting from data processing, the data should be consistently processed using a single-source processor (Hooker et al., 2001; Werdell & Bailey, 2005). Even using single-source processing, however, there will be uncertainties with in situ radiometric measurements resulting from the difficult nature of calibrating these instruments as well as deployment related issues. Hooker and Maritorena (2000) demonstrated that under ideal conditions, in situ radiometric measurement uncertainties are on the order of 3%–5%. Prerequisite ancillary information includes the date and time of collection, and the latitude and longitude of collection. Experience has suggested that useful additional information includes: wind speed and direction, wave-height, cloud

conditions, atmospheric pressure, ozone concentration, and aerosol optical thickness (Hooker et al., 1994).

2.1.1. Optically "shallow" water

Occasionally, the apparent optical depth (AOD) of the water, i.e., the depth at which penetrating light is of optical significance, exceeds the physical depth of the measurement location. When this occurs, reflectance off the bottom may affect the observed $L_{wn}(\lambda)$ (Lee et al., 1999; Spitzer & Dirks, 1987). To eliminate these cases from the validation data set, a flag is set for records where the physical depth of the measurements locations are shallower than the local AOD calculated via:

$$\mathrm{AOD} = 1.3/K_{490} \tag{1}$$

where K_{490} is the downwelling diffuse attenuation coefficient at 490 nm.

Gordon and McCluney (1975) demonstrated that, for a homogeneous water column, 90% of the water-leaving radiance signal originates from the first optical depth, or the depth at which the irradiance falls to $1/e$ of the incident irradiance value. Assuming nominal error in the K_{490} estimation, a conservative value of 1.3 is used in lieu of 1.0 to ensure that potential bottom reflectance will not influence the results. A number of the in situ data did not include either a measurement of K_{490} (e.g. above-water radiometric observations), or a recorded water depth. To ensure that this flag was consistently applied to all in situ data, the coincident satellite-derived K_{490} (Mueller & Trees, 1997; O'Reilly et al., 2000) was used for K_{490} in Eq. (1) and the NOAA ETOPO2 dataset (2001) for water depth.

2.1.2. Normalized water-leaving radiance

In situ normalized water-leaving radiances (Eq. (2)) are calculated via:

$$L_{wn} = \frac{L_w}{E_s} * F_0 \tag{2}$$

where L_w, E_s, and F_0 are spectral water-leaving radiance, surface irradiance, and the mean extraterrestrial irradiance, respectively. For E_s, theoretical clear sky surface irradiances were adopted, assuming a maritime aerosol type and an ozone concentration of 300 Dobson (e.g. Eq. (6) in Frouin et al., 1989). Theoretical surface irradiances were chosen over measured surface irradiances for several reasons. First, a number of the measured irradiances exceeded the theoretical clear sky values by more than 10% and in rare cases approached the extraterrestrial value. Second, there were occasionally large discrepancies between surface irradiance measurements calculated by extrapolation of the water column irradiance profile and those measured directly via a surface sensor (Zaneveld et al., 2001 and Fig. 1 of Werdell and Bailey, 2005). Finally, Hooker et al. (2001) found that the uncertainties in normalized water leaving radiance were consistently larger than the uncertainties in subsurface radiance and suggested that validation may be best accomplished without normalization of water leaving radiance if the data were collected very close to the satellite overpass. Unfortunately, due to the paucity of measurements made near the satellite overpasses, a larger time window is often required when defining coincidence and therefore normalization is preferred. A simple normalization to the cosine of the solar zenith angle was initially considered, however, normalization by theoretical irradiance is able to account for atmospheric transmittance as well as solar elevation and was therefore ultimately selected.

The current reprocessing of SeaWiFS data (reprocessing 5.1, July 2005) includes the bidirectional reflectance distribution function (BRDF) described by Morel and Gentili (1991, 1993, 1996) with modifications described in Gordon (2005) in the 'exact' normalization of water-leaving radiance. For consistency with the satellite product, we applied this BRDF correction to the in situ water-leaving radiance data, using the same look-up tables and approach employed by the satellite processing code.

2.2. Satellite considerations

2.2.1. Spatial considerations

The spatial resolutions of most ocean color satellites fall in the range of 300 m to 1.1 km (at nadir viewing). Given the satellite scale, the ground truth data are ideally collected in regions where the spatial variability of the geophysical parameter under consideration is relatively stable for an area several times the spatial resolution of the satellite-based instrument (Gordon et al., 1983). This accounts for possible navigation errors in the satellite data and minimizes the effect of small-scale spatial variability on the measured in situ data. Sub-pixel geophysical variability is effectively averaged by space-borne remote sensors, while in situ instrumentation does not adequately characterize this variability.

We avoid the use of SeaWiFS GAC [Global Area Coverage] resolution data, or any reduced resolution product, in the validation process. Whenever possible, we choose satellite data collected at the native resolution of the sensor. Any sub-sampling of the satellite data prior to the validation analysis adds an unquantifiable uncertainty to the results. For example, the SeaWiFS stray light correction is applied to pixels near a bright target. With the GAC resolution data set, however, the data are sub-sampled on board the spacecraft, potentially including unidentified bright targets, and thus stray light cannot be adequately corrected in the downstream processing (Barnes et al., 1995). To avoid this and other potential errors, the results presented here exclude GAC resolution data.

Satellite navigation may not be accurate to a pixel (Patt, 2002), therefore, a box of some number of pixels (e.g. 3×3, 5×5, or 7×7) is defined, centered on the location of the in situ measurement. This box allows for the generation of simple statistics, such as the mean and standard deviation, to assist in the evaluation of spatial stability, or homogeneity, at the validation point. Further, the use of a multi-pixel box increases the possibility of an in situ measurement being available for validation by increasing the chance that the satellite retrieval will have sufficient clear pixels to be useful. Following, the closest pixel to the in situ location need not be clear, so long as the valid pixels in the box meet the homogeneity requirement.

Hu et al. (2001) suggest that, due to pixelization, satellite ocean color data require aggregation (i.e. averaging) in order to meet the stated accuracy goals for the products. We compute the necessary minimum sample size for statistical reliability using:

$$N = \frac{s^2 * t^2_{\alpha(2),(n-1)}}{d^2} \quad (3)$$

where N is the number of pixels required, s^2 is the sample variance, $t_{\alpha(2),(n-1)}$ is the two-tailed critical value of Student's t test, and d is one half of the accuracy requirement (Zar, 1996). The mean, clear-water (C_a less than 0.1 mg m^{-3}, water depth greater than 1000 m) values for $L_{wn}(\lambda)$ derived from SeaWiFS are approximately 1.97 μW cm^{-2} nm^{-1} sr^{-1} and 0.28 μW cm^{-2} nm^{-1} sr^{-1} for 443 nm and 555 nm, respectively. These values agree well with theoretical values (Gordon et al., 1988; Morel & Maritorena, 2001). Given the radiometric accuracy goal of ±5% in $L_{wn}(\lambda)$ (Hooker et al., 1992), we assign values of d to be within 0.1 μW cm^{-2} nm^{-1} sr^{-1} and 0.015 μW cm^{-2} nm^{-1} sr^{-1} for 443 and 555 nm, respectively. Using an open ocean region (Hawaiian Ocean Time Series ALOHA station, Karl & Lukas, 1996), an estimate of the sample variance was made by calculating the median variance for 25 SeaWiFS scenes using a 25 × 25 pixel box where 80% of the pixels were free from cloud contamination. The use of an open ocean location reduced the possibility of geophysical variability affecting the variance estimate, and thus, we assume that the derived variances are due to the combination of inherent sensor and atmospheric correction uncertainties. The estimated sample variances were 0.00563 μW cm^{-2} nm^{-1} sr^{-1} and 0.00029 μW cm^{-2} nm^{-1} sr^{-1} for 443 nm and 555 nm, respectively.

Eq. (3) was solved iteratively on the critical Student's t value until N converged. An initial value for N of 25 pixels was used. The values of N for 443 nm and 555 nm converged at 9 and 18 pixels, respectively. Considering that not all scenes are cloud and land free, e.g. that a number of pixels will be masked as unusable, we double these values to 18 and 36 pixels. Restricting the box to odd dimensions to maintain the location of interest in the center of the box suggests a size of 7 × 7 (49 pixels). A sensitivity analysis of the validation data (results not shown) indicates that the use of a 5 × 5 box does not significantly degrade the results. We desire a small sampling box as we assume that as box size increases, geophysical variability may be introduced in violation of our requirement of homogeneity. We ultimately selected a 5 × 5 pixel box to limit the inaccuracy posed by geophysical variability, errors in navigation, and errors in the algorithms employed in the satellite retrieval process, while maintaining a reasonably small sampling area.

Careful consideration of measurement scale is critical when validating remote sensors with in situ observations, particularly that of the local geophysical phenomenon influencing those measurements. Failure to do so will often affect the interpretation of the validation results. The phenomena to be most cognizant of are those that fall between the scale of the in situ instrument and the satellite sensor. The averaging of satellite data within a box centered on the in situ location implies that the region of interest is geophysically homogeneous. As is described later, when the assumption of homogeneity could not be made, we choose not to use the data for validation purposes.

2.2.2. Coincidence determination

Satellite data are navigated to identify the pixel that corresponds with each in situ location. As the in situ data are rarely collected exactly when the satellite views their location, we assign a temporal threshold in our definition of coincidence. This time window is defined to be short enough to reduce the effects of temporal variability in the in situ data, yet sufficiently large to allow for the greatest possibility of a match. For the homogeneous water masses under consideration, we assigned a ±3-h window around the satellite overpass. We assumed that illumination is sufficient and atmospheric conditions are reasonably stable over this period (Bailey et al., 2000).

We require that each satellite record is entirely unique, having no pixels in common with any other validation record. Multiple in situ measurements per satellite box (5 × 5) are reduced to unique validation points by two methods. First, in situ measurements made at a single station are reduced to a single representative sample prior to inclusion in the validation data set (Werdell & Bailey, 2005). And second, along-track in situ measurements (i.e. multiple casts from a survey) are excluded if they fall within the bounds of the box defined for the previous validation point. The in situ data record with the smallest time difference from the satellite data record is used as the starting reference point for the along-track exclusion. It is possible, especially at high latitudes, to acquire satellite data from multiple orbits for a given in situ record. When these cases occur, the satellite orbit closest in time to the in situ measurement is selected.

2.2.3. Satellite data processing

Level 1A (instrument digital counts, L1A) satellite data are extracted to 101 × 101 pixel subsets centered on the location of the coincident in situ data. These subset L1A files are processed to Level 2 (geolocated, geophysical values, L2) using the Multi-Sensor level 1 to 2 (MSL12) processing code, version 5.2 distributed with the SeaWiFS Data Analysis System (SeaDAS) version 4.8 (Baith et al., 2001; McClain et al., 2004). Version 5.2 of MSL12 implements the atmospheric correction algorithm described in Gordon and Wang (1994) with additional corrections for BRDF (Morel & Gentili, 1991, 1993, 1996), out-of-band radiance (Wang et al., 2001; Patt et al., 2003), and near-infrared (NIR) water-leaving radiance (Patt et al., 2003). The output L2 files include: $L_{wn}(\lambda)$, the SeaWiFS default C_a product (OC4v4, O'Reilly et al., 2000), aerosol optical thickness at 865 nm, and K_{490}, as well as ancillary information such as solar and sensor zenith angles, retrieved aerosol type, and L2 processing flags.

2.2.4. Satellite data set screening

We exclude data where the viewing and solar zenith angles exceed 60° and 75°, respectively. These thresholds correspond to the those used operationally in the Level 3 binning process for the production of the SeaWiFS global products. This exclusion accounts for limitations on the reliability of

atmospheric correction algorithms at extreme viewing and solar geometries. The SeaWiFS atmospheric correction algorithm is based on plane-parallel geometry (Gordon & Wang, 1994), which ignores earth curvature. Ding and Gordon (1994) demonstrated that the plane-parallel assumption is sufficiently accurate for solar zenith angles less than about 70° and view angles up to 45°. For SeaWiFS, which tilts 20° fore or aft during normal operation, a view angle of 45° corresponds to a sensor zenith angle of approximately 56°.

The remainder of the exclusion criteria (Fig. 1) applied are based on the L2 processing flags set by the atmospheric correction algorithm (Franz, 2005). The flags are used to identify and exclude questionable pixels from the 5×5 box. Pixels are masked if any of the following flags are set: land, cloud or ice, stray light, sun glint, high top-of-atmosphere (TOA) radiance (i.e. at, or near, saturation), low $L_{wn}(555)$ (a flag used to identify cloud-shadowed pixels), or atmospheric correction failure. Experience has demonstrated that excluding pixels with these flags provides the best quality data for comparison with field data. Other flags, specific to the product under validation, are also considered for exclusion criteria. For example, SeaWiFS L2 processing provides a C_a warning flag that is set for either extreme reflectance ratio retrievals (O'Reilly et al., 1998) or for values falling outside a pre-defined range, indicating lower confidence in the C_a retrieval. We exclude such flagged pixels specifically from the C_a validation data set. It should be further noted that the flagging criteria are sensor specific, as well as a function of the atmospheric correction algorithm employed.

We require that a minimum of 50% of the pixels in the defined box be valid (i.e. unflagged) to ensure statistical confidence in the mean values retrieved. For in situ data collected in coastal waters, the valid pixel criterion is relaxed to a minimum of 50% of the non-land pixels, with an absolute minimum of 5 pixels. The arithmetic mean of the non-masked pixels in the 5×5 box, centered on the in situ location, is determined and the standard deviation calculated. To minimize the effect of outliers on the calculated mean value, especially for the case of coastal locations where the required minimum pixel count is reduced, a filtered mean value is also calculated:

$$\text{Filtered mean} = \frac{\sum_i (1.5*\sigma - \bar{X}) < X_i < (1.5*\sigma + \bar{X})}{N} \quad (4)$$

where $\bar{X}$ is the unfiltered mean value, σ is the standard deviation of the unfiltered data and N is the number of values within $\pm 1.5*\sigma$.

A revised pixel count (total minus masked pixels minus filtered pixels) is computed. The mean, median, standard deviation, filtered mean, filtered standard deviation, minimum value, maximum value, valid pixel count and filtered pixel count are stored for each satellite record.

A coefficient of variation (filtered standard deviation divided by the filtered mean, CV) is computed for the $L_{wn}(\lambda)$ retrievals

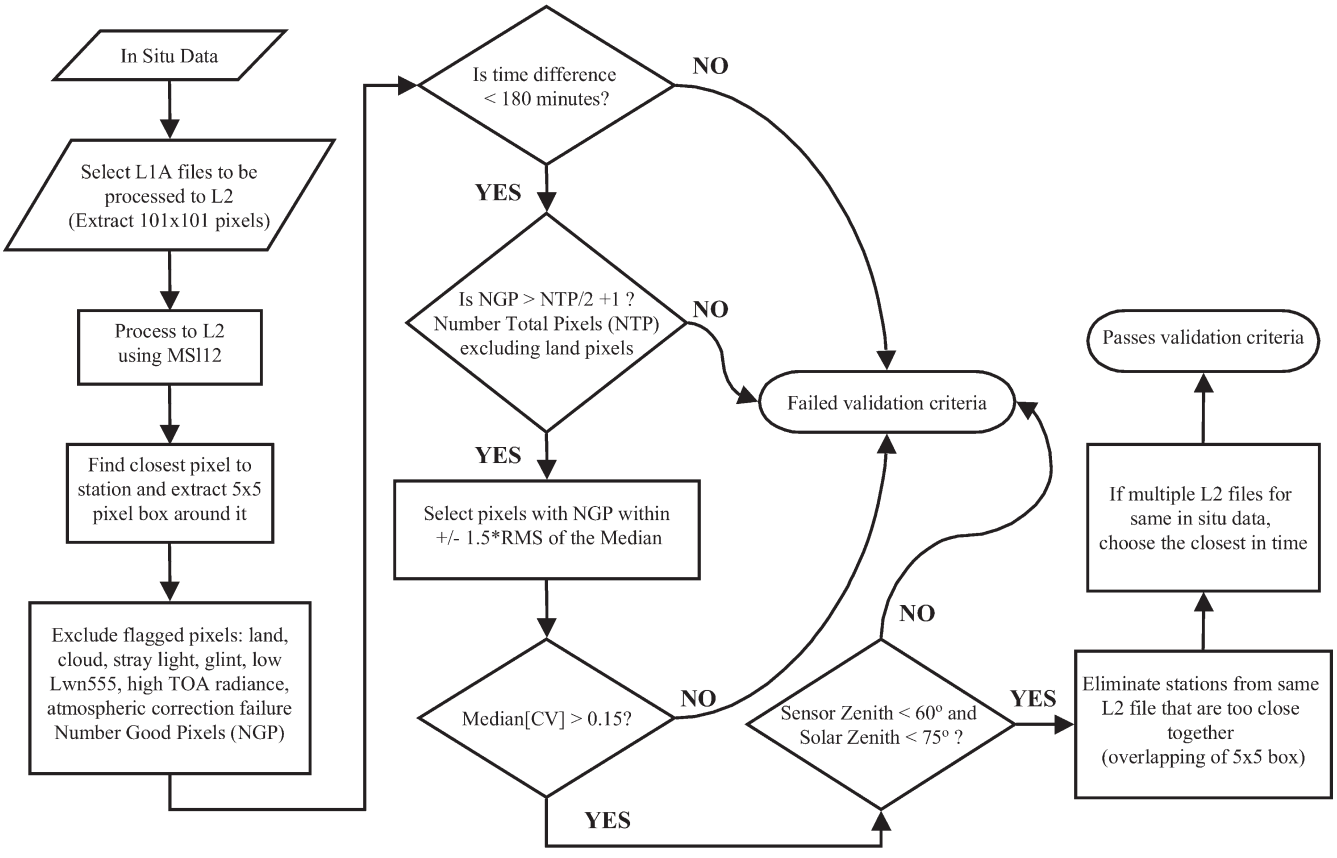

Fig. 1. Flowchart of the validation process highlighting the applied exclusion criteria.

for bands between 412 and 555 nm and for the aerosol optical thickness at 865 nm. The median CV is recorded. Finally, satellite retrievals with extreme variation between pixels in the defined box (CV greater than 0.15) are excluded. These typically represent frontal regions or other anomalies (e.g., cloud edges) which make the validation results questionable (Kilpatrick et al., 2001).

3. Validation results: SeaWiFS

We acquired 4124 and 11 059 in situ radiometric and pigment observations, respectively, with spatial and temporal coincidence with SeaWiFS data. These resulted in a maximum of 629 radiance and 1293 C_a validation records after all exclusion criteria were applied (Fig. 2). The spectral radiance sample sizes vary with wavelength because of differences in field instrumentation. Approximately 58% of the potential validation points were eliminated due to lack of sufficient valid satellite retrievals in the 5 × 5 box (i.e. below the minimum pixel threshold; 50% yielded zero retrievals), primarily due to cloud cover. Another 20% were eliminated as they exceeded the predefined allowable time window. An additional 2.5% were eliminated due to each of the apparent optical depth, coefficient of variation, geometry and the along-track distance constraints. This amounts to a return of approximately 15%.

For each individual validation point, the satellite to in situ ratio and absolute percent difference are calculated. These are summed, and the median satellite to in situ ratio, the semi-interquartile range (SIQR) for this ratio and the median absolute percent difference (MPD) are computed for each product validated (Table 1).

Table 1
Validation statistics for SeaWiFS – global

Product	Ratio (SIQR)[a]	% Difference[b]	Slope	r^2	RMSE	N
$L_{wn}(412)$	0.905 (±0.267)	24.09	1.146	0.827	0.332	480
$L_{wn}(443)$	0.915 (±0.169)	17.48	1.048	0.830	0.262	629
$L_{wn}(490)$	0.918 (±0.133)	15.10	0.979	0.821	0.207	587
$L_{wn}(510)$	0.918 (±0.124)	13.74	0.936	0.849	0.177	406
$L_{wn}(555)$	0.915 (±0.145)	16.88	0.896	0.931	0.146	629
C_a	0.998 (±0.352)	33.09	0.951	0.796	0.638	1293

As the C_a data span several orders of magnitude, a log transformation was made prior to calculation of slope and r^2 and bias statistics for this product.
[a] Median satellite to in situ ratio.
[b] Median % difference.

The SIQR provides an indication of the spread of the data and is calculated via:

$$\mathrm{SIQR} = \frac{Q_3 - Q_1}{2} \tag{5}$$

where Q_1 is the 25th percentile and Q_1 is the 75th percentile (Q_2 would be the median value). While the median ratio provides an indication of overall bias and the SIQR gives a measure of uncertainty, an alternative, unbiased statistic is required to provide additional information on how accurately the satellite retrieval agrees with in situ measurements. For this, the median absolute percent difference (MPD) is employed. MPD is calculated as the median of the individual absolute percent differences, where the absolute percent difference (PD) is calculated as:

$$\mathrm{PD}_i = 100^* \frac{|X_i - Y_i|}{Y_i} \tag{6}$$

where X is the satellite retrieved value and Y is the in situ value.

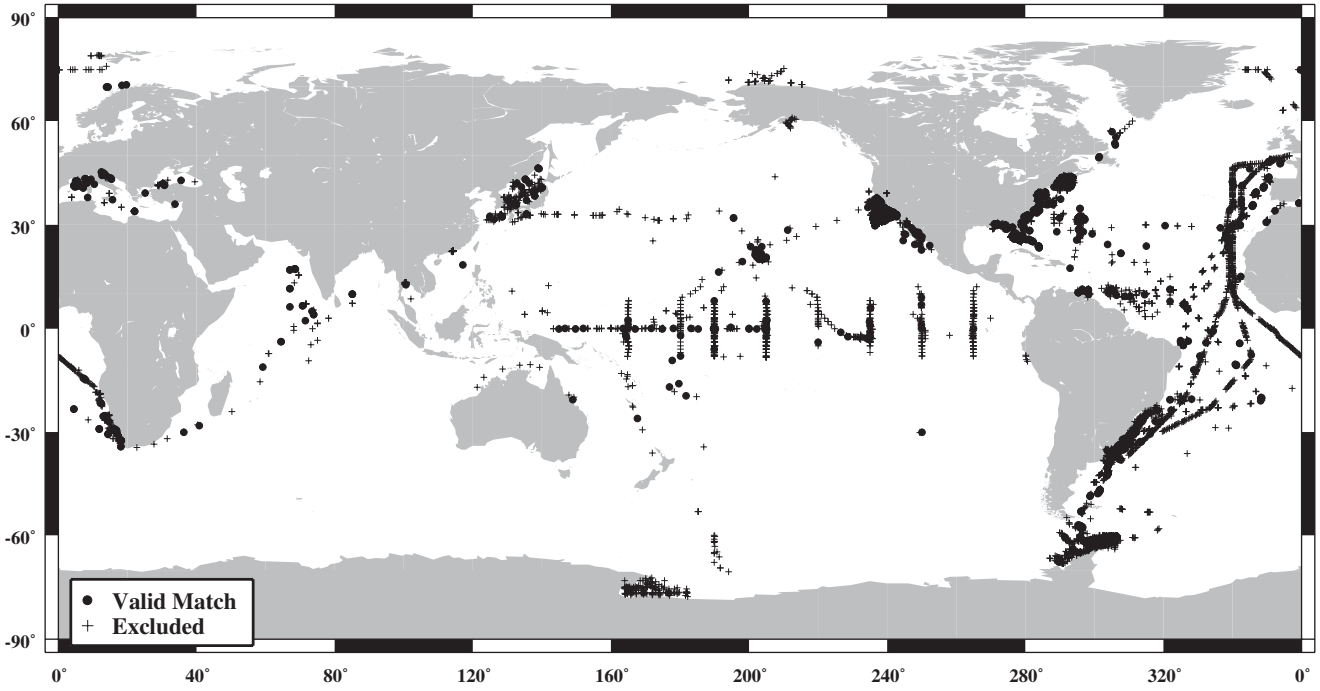

Fig. 2. Map of SeaWiFS validation data. Circles indicate a successful validation record. Crosses indicate where coincident in situ data exist, but the record failed to pass all exclusion criteria.

Scatter plots of satellite versus in situ values for each product were produced (Fig. 3), for which the slope, coefficient of determination (r^2), and root mean square error of the regression (RMSE) are listed in Table 1. The satellite-derived values were regressed against the in situ data using a reduced major axis (RMA) approach, as error may be associated with both the in situ and satellite data (Sokal & Rohlf, 1995).

The comparisons of $L_{wn}(\lambda)$ and C_a generally show reasonable agreement between SeaWiFS and in situ values. The median satellite to in situ ratios for all products are within 10% of unity, with that for C_a approaching unity. The uncertainty about the median ratio (as indicated by the SIQR) for most bands is on the order of 0.15. The $L_{wn}(412)$ ratio has notably larger SIQR at 0.267. In general, the $L_{wn}(\lambda)$ comparisons exhibit a clear bias toward under-estimation, as evident by all wavelengths having a median ratio of less than unity. The MPD for the $L_{wn}(\lambda)$ ranges between a low of 13.74% for the 510 nm band and a high of 24.09% for the 412 nm band. The C_a comparison has a MPD of 33.09%.

The median ratio, SIQR, and MPD are bulk statistics that do not provide information on how well the retrievals compare over their dynamic range. For this we examine the regression statistics (i.e. slope, r^2, and RMSE) and distribution histograms. The slopes of the regression plots are all close to unity, with the 412 nm band having a slope that deviates most significantly from unity, at 1.146. There is a trend for the slope to rotate to below unity with increasing wavelength. The r^2 for the regressions are all quite high, at or above about 0.8. The high r^2 and slopes near unity suggest that the comparisons agree over the measured dynamic range.

Table 2
Validation statistics for SeaWiFS – deep water

Product	Ratio (SIQR)	% Difference	Slope	r^2	RMSE	N
$L_{wn}(412)$	1.030 (±0.110)	10.78	1.074	0.904	0.2294	154
$L_{wn}(443)$	0.962 (±0.136)	13.34	1.066	0.854	0.2375	242
$L_{wn}(490)$	0.956 (±0.117)	12.13	0.979	0.744	0.1689	242
$L_{wn}(510)$	0.975 (±0.098)	11.08	1.240	0.478	0.1120	127
$L_{wn}(555)$	0.976 (±0.169)	16.64	0.792	0.675	0.0650	242
C_a	0.9221 (±0.239)	25.96	0.905	0.833	0.406	271

See footnotes of Table 1.

3.1. Deep water

Given the wide geographic coverage, the OBPG generates subsets of the data to study specific regional, physical, and trophic scenarios. One important case is for deep water (≥1000 m), which is representative of the majority of the global ocean. The majority of the available validation data are in complex coastal regimes, and as a result, the interpretation of the global validation results may be negatively impacted by this coastal dominance. The accuracy goals defined in Hooker et al. (1992) and (Gordon & Wang, 1994) were for clear, Morel Case-1 water (Morel & Prieur, 1977), which is approximated by our deep-water subset. The results for this deep-water subset are presented in Table 2 and Fig. 4.

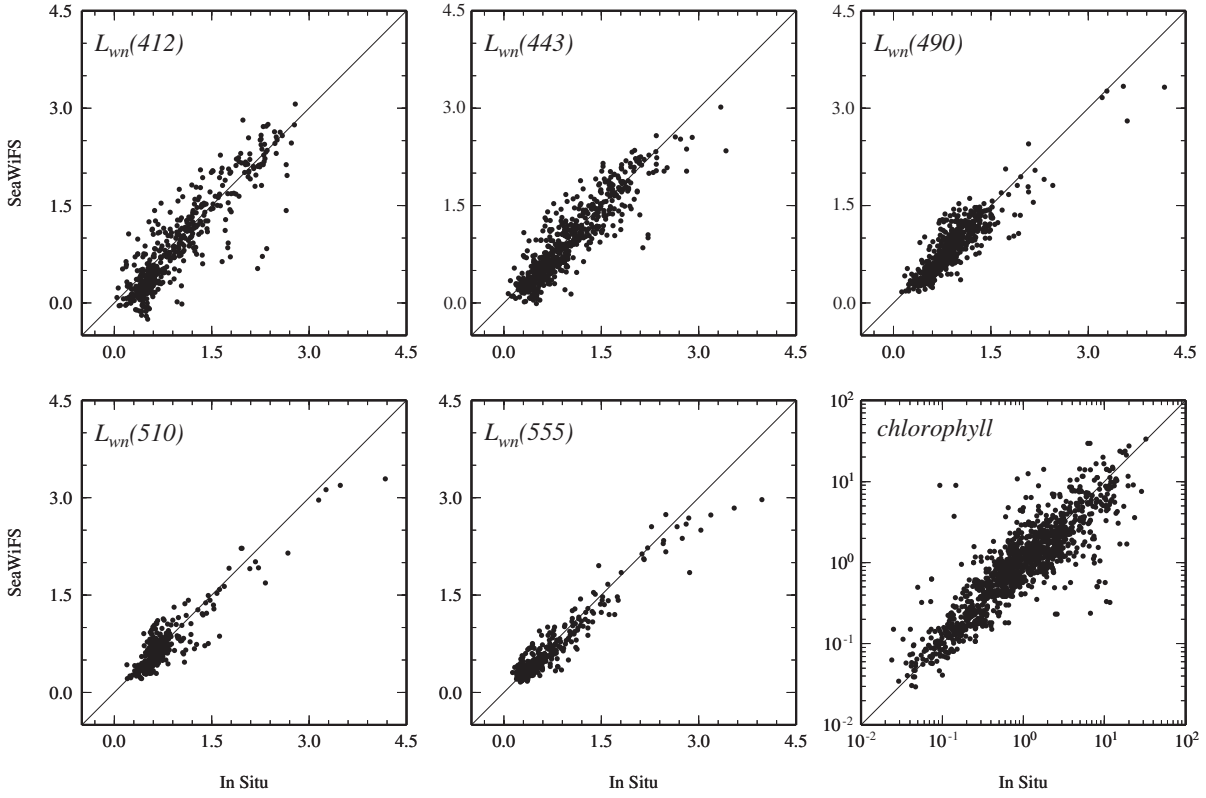

Fig. 3. Scatter plots of SeaWiFS to in situ measurements for passing all exclusion criteria.

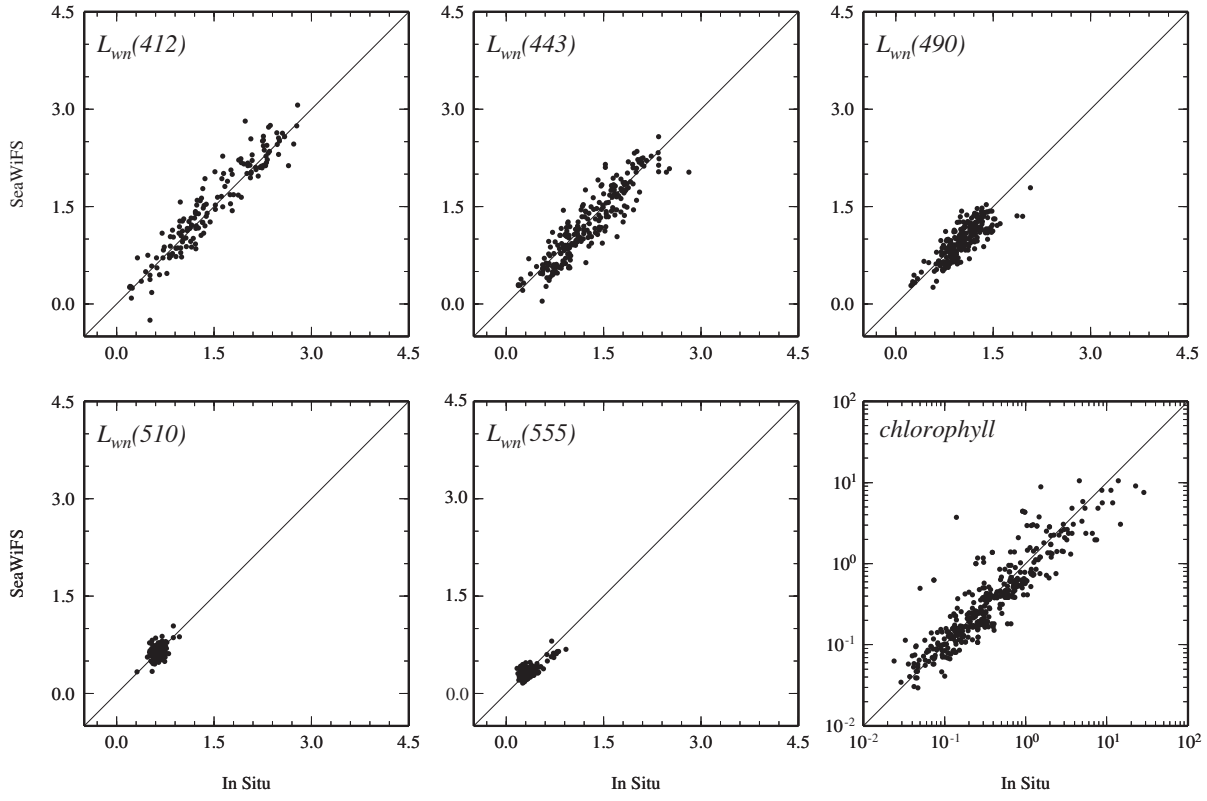

Fig. 4. Scatter plots of SeaWiFS to in situ measurements for passing all exclusion criteria for the deep water (≥1000 m) subset.

By comparison, the results of the deepwater subset show a general tendency toward better agreement with in situ over the global data set. The MPD and RMSE for all $L_{wn}(\lambda)$ and C_a are lower for the deepwater set, while the median ratios tend toward unity. In general, the slopes and r^2 for regressions in the deepwater subset are improved, though the severely reduced dynamic range for the 510 and 555 nm bands makes interpretation of the slope and r^2 less meaningful. Histograms of the percent difference for $L_{wn}(443\,nm)$ and C_a (Figs. 5 and 6) show that the deepwater subset has a slightly narrower distribution, and in the case of $L_{wn}(443\,nm)$, is more centralized about zero percent difference.

3.2. Atmospheric correction issues

A spectral bias exists in the improvement of the satellite validation results with the deepwater subset, such that the shorter wavelengths show a more dramatic improvement over the longer wavelengths. This suggests one of two likely atmospheric correction problems. The first is a failure of the black pixel assumption for the NIR and subsequent failure of the iterative NIR water-leaving radiance correction scheme (Patt et al., 2003). Such failures result in incorrect aerosol model selections, generally tending toward the identification of the tropospheric models, which have much stronger spectral dependencies than either the maritime or coastal model sets. If the true aerosol type is a maritime or coastal type, the selection of a tropospheric model will result in a spectrally biased over-subtraction of aerosol radiance, increasing with decreasing wavelength.

The second probable cause is the presence of absorbing aerosols in the coastal regions. The aerosol models used in the atmospheric correction are all non- or weakly absorbing (Gordon & Wang, 1994). Absorbing aerosols have a lower aerosol radiance at the shorter wavelengths than any of the models used in the current atmospheric correction process. The

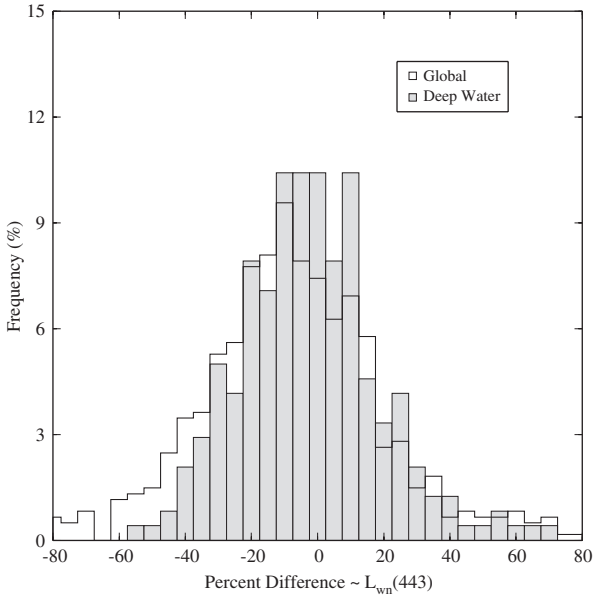

Fig. 5. $L_{wn}(443)$ percent difference histogram. Shaded histogram is for the deepwater subset, open histogram is for the full validation data set.

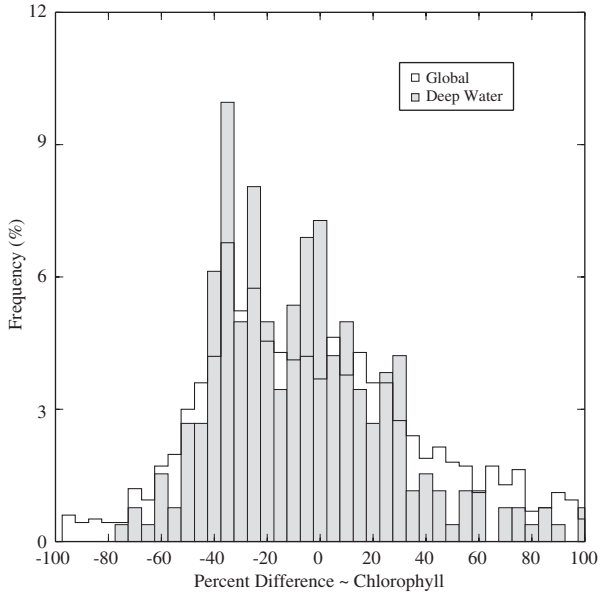

Fig. 6. C_a percent difference histogram. Shaded histogram is for the deepwater subset, open histogram is for the full validation data set.

presence of absorbing aerosols would, thus, have a similar effect to the aforementioned incorrect model selection, though likely to a more significant degree in the blue portion of the spectrum. Fig. 7 shows example in situ and SeaWiFS spectra for a case of an apparent absorbing aerosol from the Santa Barbara Channel (collected January 17, 2001). The spectra have been normalized to the radiance at 490 nm to highlight the relative effects of absorbing aerosols on the satellite retrievals.

Both highly scattering waters, which result in the black pixel assumption failure, and absorbing aerosols tend to be concentrated near the coast and, as such, are generally eliminated from the validation set with the deepwater exclusion. The impact of this can be most dramatically seen in the significant improvement of the MPD for the 412 nm band. Dust events also result in absorbing aerosols that can impact the deepwater satellite retrievals, however these tend to be more episodic than urban aerosols in the coastal zone (Prospero, 1999).

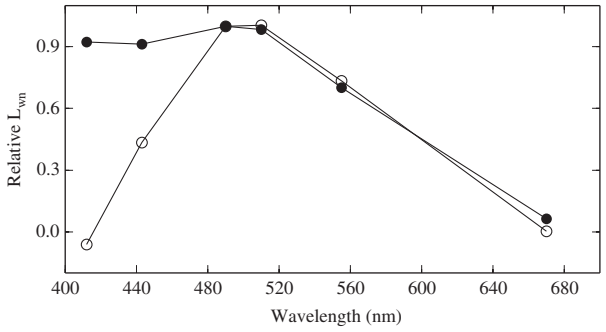

Fig. 7. Spectra for an absorbing aerosol case: (○) SeaWiFS data; (●) coincident in situ data. The spectra have been normalized to the value at 490 nm to highlight the spectral effects of absorbing aerosols on the satellite retrievals.

4. Discussion

Given the rate of return of approximately 15%, the task of validating ocean color sensors requires a dedicated in situ data collection effort to ensure that sufficient data are available to assess the satellite sensor performance on regional, global, and mission-long scales. The SeaWiFS Project Office (the precursor of the OBPG) understood this requirement early in mission planning (Hooker & McClain, 2000; Hooker et al., 1992), and initiated the development of SeaBASS to address this need. As a direct result of this foresight, the data set presented here covers a wide geographic diversity (Fig. 2) and spans the three trophic regimes–oligotrophic ($C_a \leq 0.1$ mg m^{-3}), mesotrophic ($0.1 < C_a \leq 1.0$ mg m^{-3}) and eutrophic ($C_a > 1.0$ mg m^{-3})–as defined by Antoine et al. (1996) (Fig. 8). Despite the fact that nearly 40% of the radiometric measurements are in the deepwater subset, less than 5% are from oligotrophic waters. As for chlorophyll, only slightly more than 5% of the measurements are from oligotrophic waters. The vast majority of the data, both radiometric and chlorophyll, are within the mesotrophic and eutrophic ranges.

The global SeaWiFS comparisons indicate that the remote sensor typically retrieves water-leaving radiance within 14–24% of the coincident in situ measurements. This is 3–5 times the stated goal of a 5% absolute water-leaving radiance accuracy (Hooker et al., 1992). However, as reported in Hooker and Maritorena (2000), most, if not all, of this target uncertainty can be absorbed by the uncertainty in the in situ measurements. In an effort to minimize the uncertainties with the processing of the in situ radiometric profile data, tools were developed for consistently processing data archived within SeaBASS (Werdell & Bailey, 2005). In the best case scenario, however, these

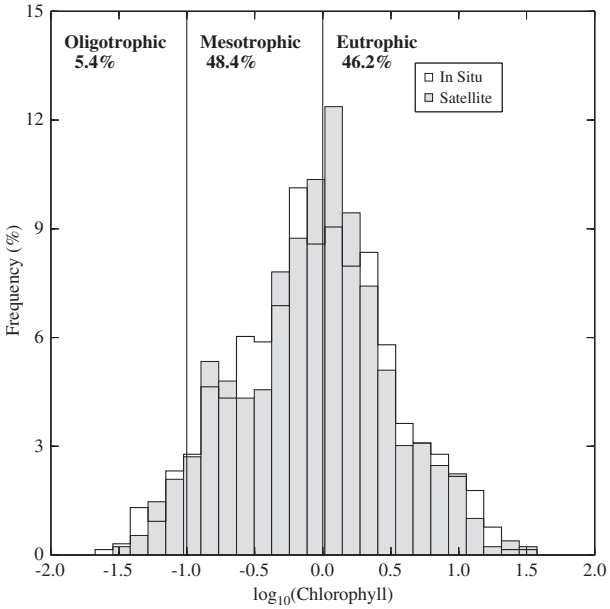

Fig. 8. C_a distribution histogram. Shaded histogram is for the SeaWiFS measurements, open histogram is for the coincident in situ measurements. Vertical lines at 0.1 mg m^{-3} (−1.0 in log) and 1.0 mg m^{-3} (0.0 in log) delineate data from the oligotrophic, mesotrophic and eutrophic regions.

uncertainties remain on the order of 3–5% (Hooker & Maritorena, 2000). If we assign a nominal (and likely conservative) 5% uncertainty to the in situ measurement, the global SeaWiFS accuracy may be considered to lie within 9–19%. Further reducing the analysis to the deepwater subset puts the SeaWiFS radiance accuracy within 6–12% for the majority of the global ocean. Note that the Level 3 products (4 km and 9 km) used in most climate research are dominated by the deepwater pixels.

A serious limitation of the current atmospheric correction algorithm is its inability to correct for absorbing aerosols. This limitation will affect the water-leaving radiance retrievals for the shorter wavelengths (e.g. 412 nm) more than those for longer wavelengths. The effect will be limited to moderate-to-low concentrations of absorbing aerosols as higher concentrations are often flagged as clouds by the MSL12 processing code. Until such time as absorbing aerosol detection and correction is successfully implemented in the atmospheric correction algorithm (e.g. Nobileau and Antoine, 2005), satellite ocean color measurements over areas with absorbing aerosols will be severely, and negatively, affected.

For both the global data set and the deepwater subset, the median satellite to in situ ratio shows a negative bias for all bands, except the deepwater $L_{wn}(412)$, which has a slight positive bias. The negative bias is consistent at approximately 9% for the global L_{wn}. For the deepwater subset, the bias is less, at about 2.5%–4%. This suggests a potential error in the estimation of the magnitude of the aerosol radiance by the atmospheric correction algorithm.

Ocean color data products are by their very nature *derived* products. For all instruments, the sensor measures top of atmosphere (TOA) radiance, and an atmospheric correction algorithm is applied to retrieve estimates of spectral water-leaving radiance. Water-leaving radiances are the principle quantities used in subsequent algorithms to produce additional derived products (e.g. the surface concentration of chlorophyll *a*). Atmospheric correction algorithms, sophisticated as they are, cannot account for all environmental conditions that may be encountered. As is evident by the improvement in the validation results for the deepwater subset, there are conditions, such as absorbing aerosols and highly scattering waters, that are not adequately addressed by the current atmospheric correction algorithms. Therefore, sensor characteristics and algorithm limitations need to be understood and environmental conditions need to be known, so that validation results can be interpreted correctly. An understanding of the accuracy of all the algorithms involved in the derivation of a particular satellite-derived product is essential in the validation process (Wang et al., 2005). When drawing conclusions about validation results for secondary derived products, it must be stressed that the satellite product accuracy will never exceed the accuracy of the algorithm from which it is derived. In the case of the SeaWiFS C_a algorithm (OC4v4, O'Reilly et al., 2000), the stated algorithm RMSE is 0.222 in log units. For a C_a of 1 mg m^{-2}, this is equivalent to an RMSE of 0.599 in normal space. The RMSE for the global SeaWiFS C_a validation set is 0.6382, very similar to the RMSE for the algorithm itself.

Validation of remotely sensed ocean color data products is a complicated task, one requiring a large in situ data set that both temporally spans the lifetime of the satellite mission and is geographically global in scope. In this paper, we have outlined an approach to validation that is operationally applied by the OBPG to most ocean color sensors. There are a few key recommendations we make from our experience applying the described method to SeaWiFS data:

(1) Use a consistently processed in situ data set
(2) Eliminate suspect in situ data (e.g. from optically shallow waters) from the validation set
(3) Use a narrow time window for determining coincidence (i.e. no more than ±3 h) between in situ and satellite data records
(4) Use native resolution satellite products (i.e. avoid sub-sampled data)
(5) Use the mean of a 5 × 5 pixel box centered on the in situ location
(6) Appropriately mask satellite pixels on the L2 flags
(7) Use a homogeneity test (e.g. CV) to minimize the impact of geophysical variability in the 5 × 5 pixel box on the satellite measurement mean

Following these recommendations will aid in the analysis of the resulting validation results by minimizing the systemic uncertainties.

Acknowledgments

This work was funded by the NASA Earth Observing System (EOS)/MODIS and NASA Ocean Biogeochemistry Programs. The authors are grateful to C. McClain, G. Feldman, G. Fargion, and all OBPG Staff members for their support and assistance in this effort. We also wish to thank all of the investigators who contributed their data to SeaBASS making such a comprehensive validation effort possible. The utility and success of SeaBASS and our validation activity relies on the tremendous efforts put forth by the data contributors.

References

Antoine, D., Andre, J. M., & Morel, A. (1996). Oceanic primary production: 2. Estimation at global scale from satellite (CZCS) chlorophyll. *Global Biogeochemical Cycles*, *10*, 57−70.

Bailey, S. W., McClain, C. R., Werdell, P. J., & Schieber, B. D. (2000). Normalized water-leaving radiance and chlorophyll *a* match-up analyses. *NASA Tech. Memo.*, *vol. 206892*. Greenbelt, MD: National Aeronautics and Space Administration, Goddard Space Flight Center.

Baith, K., Lindsay, R., Fu, G., & McClain, C. R. (2001). Data analysis system developed for ocean color satellite sensors. *EOS Transactions*, *82*.

Barbini, R., Colal, F., Fantoni, R., Fiorani, L., Okladnikov, I. G., & Palucci, A. (2005). Comparison of SeaWiFS, MODIS-Terra and MODIS-Aqua in the Southern Ocean. *International Journal of Remote Sensing*, *26*(11), 2471−2478.

Barnes, R. A., Holmes, A. W., & Esaias, W. E. (1995). Stray light in the SeaWiFS radiometer. *NASA Tech. Memo.*, *vol. 104566*. Greenbelt, MD: National Aeronautics and Space Administration, Goddard Space Flight Center.

Barnes Jr., R. A., R. E. E., Schmidt, G. M., Patt, F. S., & McClain, C. R. (2001). Calibration of SeaWiFS: I. Direct techniques. *Applied Optics*, *40*(36), 6682−6700.

Barnes, R. A., Clark, D. K., Esaias, W. E., Fargion, G. S., Feldman, G. C., & McClain, C. R. (2003). Development of a consistent multi-sensor global ocean color time series. *International Journal of Remote Sensing*, *24*, 4047−4064.

Behrenfeld, M. J., & Falkowski, P. G. (1997a). A consumer's guide to phytoplankton primary productivity models. *Limnology and Oceanography*, *42*(7), 1479−1491.

Behrenfeld, M. J., & Falkowski, P. G. (1997b). Photosynthetic rates derived from satellite-based chlorophyll concentration. *Limnology and Oceanography*, *42*(1), 1−20.

Ding, K., & Gordon, H. R. (1994). Atmospheric correction of ocean color sensors: Effects of the Earth's curvature. *Applied Optics*, *33*(30), 7096−7106.

Donlon, C. J., Minnett, P. J., Gentemann, C., Nightingale, T. J., Barton, I. J., Ward, B., et al. (2002). Toward improved validation of satellite sea surface skin temperature measurements for climate research. *Journal of Climate*, *15*, 353−369.

D'Ortenzio, F., Marullo, S., Ragni, M., d'Alcala, M. R., & Santoleri, R. (2002). Validation of empirical SeaWiFS algorithms for chlorophyll-*a* retrieval in the Mediterranean Sea: A case study for oligotrophic seas. *Remote Sensing of Environment*, *82*, 79−94.

Fargion, G. S., Franz, B. A., Kwiatkowska, E. J., Pietras, C. M., Bailey, S. W., Gales, J., et al. (2004). SIMBIOS program in support of ocean color missions: 1997–2003. In R. J. Frouin, G. D. Gilbert, & D. Pan (Eds.), *Ocean remote sensing and imaging: II. Proceedings SPIE*, *vol. 5155* (pp. 49–60). The Society of Photo-Optical Instrumentation Engineers.

Franz, B. A. (2005). Msl12 user's guide. URL http://oceancolor.gsfc.nasa.gov/DOCS/MSL12/

Franz, B. A., Werdell, P. J., Meister, G., Bailey, S. W., Eplee, R. E., Feldman, G. C., et al. (2005). The continuity of ocean color measurements from SeaWiFS to MODIS. In J. J. Butler (Ed.), *Earth observing systems: X. Proceedings SPIE*, *vol. 5882* (pp. 49–60). The International Society for Optical Engineering.

Froidefond, J. -M., Lavender, S., Laborde, P., Herbland, A., & Lafon, V. (2002). SeaWiFS data interpretation in a coastal area in the Bay of Biscay. *International Journal of Remote Sensing*, *23*(5), 881−904.

Frouin, R., Ligner, D. W., & Gautier, C. (1989). A simple analytical formula to compute clear sky total and photosynthetically available solar irradiance at the ocean surface. *Journal of Geophysical Research*, *94*, 9731−9742.

Gohin, F., Druon, J. N., & Lampert, L. (2002). A five channel chlorophyll concentration algorithm applied to SeaWiFS data processed by SeaDAS in coastal waters. *International Journal of Remote Sensing*, *23*(8), 1639−1661.

Gordon, H. R. (1998). In-orbit calibration strategy for ocean color sensors. *Remote Sensing of Environment*, *63*, 265−278.

Gordon, H. R. (2005). Normalized water-leaving radiance: Revisiting the influence of surface roughness. *Applied Optics*, *44*(2), 241−248.

Gordon, H. R., & McCluney, W. R. (1975). Estimation of the depth of sunlight penetration in the sea for remote sensing. *Applied Optics*, *14*(2), 413−416.

Gordon, H. R., & Wang, M. (1994). Retrieval of water-leaving radiance and aerosol optical thickness over the oceans with SeaWiFS: A preliminary algorithm. *Applied Optics*, *33*, 443−452.

Gordon, H. R., Clark, D. K., Brown, J. W., Brown, O. B., Evans, R. H., & Broenkow, W. W. (1983). Phytoplankton pigment concentrations in the Middle Atlantic Bight: Comparison of ship determinations and CZCS estimates. *Applied Optics*, *22*(1), 20−36.

Gordon, H. R., Brown, O. B., Evans, R. H., Brown, J. W., Smith, R. C., Baker, K. S., et al. (1988). A semianalytic radiance model of ocean color. *Journal of Geophysical Research*, *93*(D9), 10909−10924.

He, M. -X., Liu, Z. -S., Du, K. -P., Li, L. -P., Chen, R., Carder, K. L., et al. (2000). Retrieval of chlorophyll from remote-sensing reflectance in the China seas. *Applied Optics*, *39*(15), 2467−2474.

Hooker, S. B., & Maritorena, S. (2000). An evaluation of oceanographic radiometers and deployment methodologies. *Journal of Atmospheric and Oceanic Technology*, *17*(6), 811−830.

Hooker, S. B., & McClain, C. R. (2000). The calibration and validation of SeaWiFS data. *Progress in Oceanography*, *45*, 427−465.

Hooker, S. B., Esaias, W. E., Feldman, G. C., Gregg, W. W., & McClain, C. R. (1992). An overview of SeaWiFS and ocean color. *NASA Tech. Memo.*, *vol. 104566*. National Aeronautics and Space Administration, Goddard Space Flight CenterGreenbelt, MD.

Hooker, S. B., McClain, C. R., Firestone, J. K., Westphal, T. L., Yeh, E. -N., & Ge, Y. (1994). The SeaWiFS bio-optical archive and storage system (SeaBASS): Part 1. *NASA Tech. Memo.*, *vol. 104566*. Greenbelt, MD: National Aeronautics and Space Administration, Goddard Space Flight Center.

Hooker, S. B., Zibordi, G., Berthon, J. -F., D'Alimonte, D., Maritorena, S., McLean, S., et al. (2001). Results of the second SeaWiFS data analysis round Robin, March 2000 (DARR-00). *NASA Tech. Memo.*, *vol. 206892*. Greenbelt, MD: National Aeronautics and Space Administration, Goddard Space Flight Center.

Hovis, W. A., Clark, D. K., Anderson, F., Austin, R. W., Wilson, W. H., Baker, E. T., et al. (1980). NIMBUS-7 coastal zone color scanner–system description and initial imagery. *Science*, *210*, 60−63.

Hu, C., Carder, K. L., & Muller-Karger, F. E. (2001). How precise are SeaWiFS ocean color estimates? Implications of digitization-noise errors. *Remote Sensing of Environment*, *76*, 239−249.

Iwasaki, N., Kajii, M., Tange, Y., Miyachi, Y., Tanaka, T., Sato, R., et al. (1992). Status of ADEOS mission sensors. *ACTA Astronautica*, *28*, 139−146.

Karl, D. M., & Lukas, R. (1996). The Hawaiian Ocean Time-series (HOT) program: Background, rationale, and field implementation. *Deep Sea Research II*, *43*, 129−156.

Kilpatrick, K. A., Podesta, G. P., & Evans, R. (2001). Overview of the NOAA/NASA advanced very high resolution radiometer Pathfinder algorithm for sea surface temperature and associated matchup database. *Journal of Geophysical Research*, *106*(C5), 9179−9198.

Kuring, N., Lewis, M. R., Platt, T., & O'Reilly, J. E. (1990). Satellite-derived estimates of primary production on the northwest Atlantic continental shelf. *Continental Shelf Research*, *10*(5), 461−484.

Lee, Z., Carder, K. L., Mobley, C. D., Steward, R. G., & Patch, J. S. (1999). Hyperspectral remote sensing for shallow waters: 2. Deriving bottom depths and water properties by optimization. *Applied Optics*, *38*(18), 3831−3843.

Longhurst, A., Sathyendranath, S., Platt, T., & Caverhill, C. (1995). An estimate of global primary production in the ocean from satellite radiometer data. *Journal of Plankton Research*, *17*(6), 1245−1271.

McClain, C. R., Esaias, W., Feldman, G. C., Frouin, R., Gregg, W., & Hooker, S. B. (2002). The proposal for the NASA Sensor Intercalibration and Merger for Biological and Interdisciplinary Oceanic Studies (SIMBIOS) Program. *NASA Tech. Memo.*, *vol. 2002-210008*. Greenbelt, MD: National Aeronautics and Space Administration, Goddard Space Flight Center.

McClain, C. R., Feldman, G. C., & Hooker, S. B. (2004). An overview of the SeaWiFS project and strategies for producing a climate research quality global ocean bio-optical time series. *Deep-Sea Research II*, *51*, 5−42.

Morel, A., & Gentili, B. (1991). Diffuse reflectance of oceanic waters: Its dependence on sun angle as influenced by the molecular scattering contribution. *Applied Optics*, *30*, 4427−4438.

Morel, A., & Gentili, B. (1993). Diffuse reflectance of oceanic waters: II. Bidirectional aspects. *Applied Optics*, *32*, 6864−6879.

Morel, A., & Gentili, B. (1996). Diffuse reflectance of oceanic waters: III. Implication of bidirectionality for the remote-sensing problem. *Applied Optics*, *35*, 4850−4862.

Morel, A., & Maritorena, S. (2001). Bio-optical properties of oceanic waters: A reappraisal. *Journal of Geophysical Research*, *106*, 7163−7180.

Morel, A., & Prieur, L. (1977). Analysis of variations in ocean color. *Limnology and Oceanography*, *22*, 709−722.

Mueller, J. L., Austin, R. W., Morel, A., Fargion, G. S., & McClain, C. R. (2003). Ocean optics protocols for satellite ocean color sensor validation: Revision 4. Introduction, background and conventions. *NASA Tech. Memo.*, *vol. 2003-21621*. Greenbelt, MD: National Aeronautics and Space Administration, Goddard Space Flight Center.

Mueller, J. L., Pietras, C., Hooker, S. B., Austin, R. W., Miller, M., Knobelspiesse, K. D., et al. (2003). Ocean optics protocols for satellite ocean color sensor validation: Revision 4. Instrument specifications,

characterization and calibration. *NASA Tech. Memo.*, *vol. 2003-21621*. Greenbelt, MD: National Aeronautics and Space Administration, Goddard Space Flight Center.

Mueller, J. L., & Trees, C. C. (1997). Revised SeaWiFS prelaunch algorithm for the diffuse attenuation coefficient K(490). *NASA Tech. Memo.*, *vol. 104566*. Goddard Space Flight Center, Greenbelt, MD: National Aeronautics and Space Administration.

NOAA ETOPO2 dataset, 2001. U.S. Department of Commerce, National Oceanic and Atmospheric Administration, National Geophysical Data Center, 2-minute Gridded Global Relief Data (ETOPO2). http://www.ngdc.noaa.gov/mgg/fliers/01mgg04.html

Nobileau, D., & Antoine, D. (2005). Detection of blue-absorbing aerosols using near infrared and visible (ocean color) remote sensing observations. *Remote Sensing of Environment*, *95*, 368−387.

O'Reilly, J., et al. (2000). SeaWiFS postlaunch calibration and validation analyses: Part 3. *NASA Tech. Memo.*, *vol. 206892*. Greenbelt, MD: National Aeronautics and Space Administration, Goddard Space Flight Center.

O'Reilly, J. E., Maritorena, S., Mitchell, B. G., Siegel, D. A., Carder, K. L., Garver, S. A., et al. (1998). Ocean color chlorophyll algorithms for SeaWiFS. *Journal of Geophysical Research*, *103*, 24937−24953.

Patt, F. S. (2002). Navigation algorithms for the SeaWiFS mission. *NASA Tech. Memo.*, *vol. 206892*. Greenbelt, MD: National Aeronautics and Space Administration, Goddard Space Flight Center.

Patt, F. S., Barnes, R. A., R.E. Eplee, J., Franz, B. A., Robinson, W. D., Feldman, G. C., et al. (2003). Algorithm updates for the fourth SeaWiFS data reprocessing. *NASA Tech. Memo.*, *vol. 206892*. Greenbelt, MD: National Aeronautics and Space Administration, Goddard Space Flight Center.

Pinkerton, M. H., & Aiken, J. (1999). Calibration and validation of remotely sensed observations of ocean color from a moored data buoy. *Journal of Atmospheric and Oceanic Technology*, *16*, 915−923.

Prasad, K. S., & Haedrich, R. L. (1994). Satellite-derived primary production estimates from the Grand Banks: Comparison to other ocean regimes. *Continental Shelf Research*, *14*(15), 1677−1687.

Prospero, J. M. (1999). Long-term measurements of the transport of African mineral dust to the southeastern United States: Implications for regional air quality. *Journal of Geophysical Research*, *104*(D13), 15917−15928.

Rast, M., & Bezy, J. L. (1999). The ESA Medium Resolution Imaging Spectrometer MERIS: A review of the instrument and its mission. *International Journal of Remote Sensing*, *20*(9), 1681−1702.

Salomonson, V. V., Barnes, W. L., Maymon, P. W., Montgomery, H. E., & Ostrow, H. (1989). MODIS: Advanced facility instrument for studies of the Earth as a system. *IEEE Geoscience and Remote Sensing Society Newsletter*, *27*, 145−152.

Smyth, T. J., Groom, S. B., Cummings, D. G., & Llewellyn, C. A. (2002). Comparison of SeaWiFS bio-optical chlorophyll-a algorithms within the OMEXII programme. *International Journal of Remote Sensing*, *23*(11), 2321−2326.

Sokal, R. R., & Rohlf, F. J. (1995). *Biometry: The principles and practice of statistics in biological research.* (3rd edition). W.H. Freeman.

Spitzer, D., & Dirks, R. W. J. (1987). Bottom influence of the reflectance of the sea. *International Journal of Remote Sensing*, *8*(3), 279−290.

Wang, M., Franz, B. A., Barnes, R. A., & McClain, C. R. (2001). Effects of spectral bandpass on SeaWiFS-retrieved near-surface optical properties of the ocean. *Applied Optics*, *40*(3), 343−348.

Wang, P., Boss, E. S., & Roesler, C. (2005). Uncertainties of inherent optical properties obtained from semianalytical inversions of ocean color. *Applied Optics*, *44*(19), 4074−4085.

Welsch, C., Swenson, H., Cota, S. A., DeLuccia, F., Haas, J. M., Schueler, C., et al. (2001). VIIRS (Visible Infrared Imager Radiometer Suite): A next generation operational environmental sensor for NPOESS. *IGARRS*.

Werdell, P. J., & Bailey, S. W. (2005). An improved bio-optical data set for ocean color algorithm development and satellite data product validation. *Remote Sensing of Environment*, *98*, 122−140.

Werdell, P. J., Bailey, S. W., Fargion, G., Pietras, C., Knobelspiesse, K., Feldman, G., et al. (2003). Unique data repository facilitates ocean color satellite validation. *EOS Transactions*, *84*(38).

Zaneveld, J. R. V., Boss, E., & Barnard, A. (2001). Influence of surface waves on measured and modeled irradiance profiles. *Applied Optics*, *40*(9), 1442−1449.

Zar, J. H. (1996). *Biostatistical analysis* (3rd edition). Prentice Hall.